

# Cross-cultural Similarities and Differences in Music Mood Perception

Jin Ha Lee<sup>1</sup> & Xiao Hu<sup>2</sup>

<sup>1</sup> University of Washington, Information School

<sup>2</sup> University of Hong Kong

## Abstract

Prior research suggests that listeners from different cultural backgrounds appreciate music differently. Although music mood/emotion is an important part of music seeking and appreciation, few cross-cultural music information retrieval (MIR) studies focus on music mood. Moreover, existing studies on cross-cultural music perception often only compare listeners from two cultures, in most cases, Western vs. Non-western cultures. In order to fill these gaps, this study compares music mood perceptions of listeners from three distinct cultures: American, Korean, and Chinese. Our findings reveal that the perceptions of the three cultural groups are generally different, but in many aspects, Korean listeners are situated in between listeners from the two other cultures. This paper describes the comparison of the three cultural groups from the perspectives of mood perceptions, musical (stimuli) characteristics, and listeners' (subjects) characteristics. The findings of this study have implications for the design of cross-cultural and global MIR systems.

**Keywords:** music information retrieval, mood, cross-cultural study

**Citation:** Lee, J. H., & Hu, X. (2014). Cross-cultural Similarities and Differences in Music Mood Perception. In *iConference 2014 Proceedings* (p. 259–269). doi:10.9776/14081

**Copyright:** Copyright is held by the authors.

**Contact:** jinhalee@uw.edu, xiaoxhu@hku.hk

## 1 Introduction

Over the past decade, much of the music information retrieval (MIR) research has been focusing on popular or classical music from Western cultures (Serra et al., 2013). As a result, there is a lack of research that can help us tackle the challenges in developing technologies that reflect musical diversity in different cultural contexts. At present, the issue of cultural diversity is ever more important in the field of MIR due to the fact that music is increasingly being appreciated across cultural boundaries through globally accessible websites such as YouTube as well as various social media. This study aims to contribute to improving our understanding of multicultural issues in MIR, specifically how people from different cultural backgrounds perceive the mood of music. In MIR, music mood has recently emerged as a potential metadata element or feature for better organization of and access to music (Lee et al., 2012). Yet, our general understanding of how real users perceive music mood is still lacking, and especially so with regard to users from multiple cultural backgrounds.

This study attempts to fill two significant gaps in current research. First, although the number of cross-cultural MIR studies has gradually been increasing, few studies explore music mood/emotion which is an important part of music appreciation. Second, many cross-cultural studies only compare two cultures, usually Western vs. non-Western cultures (e.g., Indian and European listeners in Gregory and Varney (1996), New York and Hong Kong music users in Nettamo et al. (2006), Korean and North American users in Lee et al. (2005), and Western and African listeners in Eerola et al. (2006)). Few studies explore how music information behaviors of users from multiple non-Western cultural groups compare to one another, and to a Western cultural group. We selected three cultural groups for comparison (American, Korean, and Chinese) so that we are not only comparing Western vs. non-Western users, but also two non-Western user groups.

Korean and Chinese user groups were chosen because of their unique cultural distinctions and connections. As Korea and China are geographically close to each other, there have been many interactions between the two countries in their long history of thousands of years. Therefore Korean and Chinese cultures have developed common elements and similar traditions. However, over the last century South Korea has been heavily influenced by Western cultures, specifically, American culture, primarily due to America's involvement in Korea following World War II. For decades, Korean people have been actively listening to Western music, and American pop songs are well-received and quite popular in Korea. China, on the other hand, has remained much more isolated from the influences of Western culture until fairly recently, about two decades ago. We wanted to understand how similar or different these two non-Western user groups are, compared to the American user group. A comparison between more than two cultural groups will bring us one step further towards building cross-cultural and global MIR systems.

## 2 Related Work

Existing comparative studies of Western and non-Western music listeners in music psychology tend to focus on aspects other than music mood, for example, perception of complexity (Eerola et al., 2006), the role of music in everyday life (Rana & North, 2007), functions of music and music preference (Schäfer et al., 2008), and so on. However, a small number of studies do examine how music mood judgments can transcend cultural boundaries. In some studies, listeners were asked to identify the mood of music from different cultures (e.g., Balkwill and Thompson (1999), Fritz et al. (2009)). These studies discovered that some basic emotions such as happiness, sadness, anger, or fear are recognizable by listeners across cultures. More specifically, Balkwill and Thompson (1999) discovered that Western listeners were able to identify intended emotion such as joy, sadness, and anger in Hindustani raga. Fritz et al. (2009) found that native African listeners could identify happy, sad, and scared/fearful emotions in Western music. In other studies, listeners from different cultures were asked to describe the mood of music from single or multiple cultures (e.g., Wong et al. (2009), Gregory and Varney (1996), Hu and Lee (2012)). They found many subtle differences in how mood is perceived by listeners from multiple cultures, and that cultural tradition and background, in addition to the inherent qualities of the music, affect listeners' mood judgments. Wong et al. (2009) revealed the in-culture bias of Indian and Western listeners when judging the tension in music. Gregory and Varney (1996) found that Indian and European listeners show a number of subtle differences in their mood descriptors and cultural tradition was more important than inherent qualities of music when listeners determined the music mood. Hu and Lee (2012) also found that mood judgments do differ between American and Chinese listeners.

This study builds upon these works, in particular, our previous work comparing music mood perceptions of American and Chinese listeners (2012). Some important findings of our previous work include that listeners from different cultural backgrounds tended to give different mood labels when the mood of a song was not very obvious. Listeners agree more with other listeners from the same cultural group than with those from the other cultural group. American and Chinese listeners had different opinions on all the five music genres in consideration (i.e., Dance, Easy listening, Pop, Rock, and Other), and listeners' gender and age did not have as strong an influence as cultural background on their judgments of music mood. It may not be surprising that many differences exist between listeners from American and Chinese cultures, since the two are often recognized as representing the contrasting Western and Eastern cultures. The case of Korea is a bit more intriguing, as Korean culture shares similar roots with Chinese culture, but is also strongly influenced by American culture.

To summarize, this study distinguishes itself from related cross-cultural MIR studies in that: 1) it considers three cultural groups, with one of them highly influenced by the other two; 2) explicitly measures and compares the distance among the three cultural groups; and 3) the songs (stimuli) used in this study cover a range of musical characteristics (i.e., instrumental vs. vocal, different genres).

### 3 Research Question and Method

#### 3.1 Research Question

The overarching question in this study is how similarly or differently listeners from American, Korean, and Chinese cultural backgrounds perceive music mood. We attempt to answer this question from the three following perspectives: mood judgment distributions and agreement levels, music (stimuli) characteristics, and listeners' (subjects) characteristics.

#### 3.2 Study Design

Listeners from the three cultural groups were recruited to answer an online music listening survey consisting of thirty 30-second music clips. The clips were selected from Western songs: half from the APM (Associate Production Music) library<sup>1</sup> (instrumental music in a range of different genres), and the other half from the USPOP collection (Ellis et al., 2003) consisting of Pop and Rock songs with lyrics. All the songs had previously been evaluated by three to five listeners whose cultural backgrounds ranged from Europe, North America, and Asia. In order to avoid selecting the “obvious” songs with highest agreement, we selected songs that had greater disagreement among the listeners. The music pieces from the APM library were instrumental, and therefore we balanced our test dataset by drawing the other half of the 30 pieces from the USPOP collection (Ellis et al., 2003) ensuring that they all had vocal components. More detailed description on how the songs are selected can be found in Hu and Lee (2012). In responding to the survey, participants listened to each of the clips and selected one out of five mood clusters that best describes the mood expressed by the clip. Each mood cluster contains several mood terms that collectively better represent a particular mood as opposed to a single mood term. If none of the mood clusters were deemed appropriate, the listener could choose the “Other” option and specify the mood in their own vocabulary.

The five mood clusters were adopted from the Audio Music Mood Classification task in the community-based evaluation framework MIREX (Music Information Retrieval Evaluation eXchange) (Hu et al., 2008) (reprinted in Table 1). MIREX is the venue for comparing state-of-the-art algorithms and systems that are relevant for music information retrieval. The Audio Music Mood Classification task was first included in 2007 and a number of relevant algorithms were evaluated every year since then. The five mood clusters used in MIREX Audio Music Mood Classification task were derived by conducting a hierarchical clustering on a co-occurrence matrix of music mood labels from the All Music Guide<sup>2</sup>. They are commonly used in a number of MIR studies on music mood evaluation and comparison such as Bischoff et al. (2009), Laurier et al. (2009), Dang and Shirai (2009), Panda and Paiva (2012), to name a few.

<b>Cluster 1 (C_1)</b>	passionate, rousing, confident, boisterous, rowdy
<b>Cluster 2 (C_2)</b>	rollicking, cheerful, fun, sweet, amiable/ good natured
<b>Cluster 3 (C_3)</b>	literate, poignant, wistful, bittersweet, autumnal, brooding
<b>Cluster 4 (C_4)</b>	humorous, silly, campy, quirky, whimsical, witty, wry
<b>Cluster 5 (C_5)</b>	aggressive, fiery, tense/anxious, intense, volatile, visceral

Table 1: Five mood clusters used in MIREX

The listeners were recruited via mailing lists for university students. Listeners in each group identified themselves as “American” raised in the United States, “Korean” raised in South Korea, or “Chinese” raised in Mainland China. The survey was translated into Korean and Chinese (simplified characters) for respective

<sup>1</sup> Accessible at: <http://www.apmmusic.com/apm-libraries>

<sup>2</sup> Accessible at: <http://www.allmusic.com>

user groups in order to help them better understand the meanings of the mood labels. Table 2 shows the translation of the mood terms.

C_1	<b>Korean:</b> 열정적인, 흥분시키는, 자신감 있는, 활기가 넘치는, 소란스러운 <b>Chinese:</b> 有激情的, 使人兴奋的, 自信的, 热闹的, 吵闹的
C_2	<b>Korean:</b> 흥겨운, 쾌활한, 즐거운, 달콤한, 상냥한/온화한 <b>Chinese:</b> 愉快的, 快乐的, 有趣的, 甜美的, 可亲的/温厚的
C_3	<b>Korean:</b> 세련된, 비통한, 아쉬운, 달콤쌉쌀한, 가을느낌의, 음울한 <b>Chinese:</b> 优美的, 凄美的, 怅惘的, 苦乐参半的, 悲秋的, 忧郁的
C_4	<b>Korean:</b> 익살스러운, 우스꽝스러운, 과장된, 별난, 기발한, 재치있는, 비꼬는 <b>Chinese:</b> 幽默的, 傻乎乎的, 哗众取宠的, 古怪的, 异想天开的, 诙谐的, 讽刺的
C_5	<b>Korean:</b> 공격적인, 불같은, 긴장된/불안한, 강렬한, 변덕스러운, 본능적인 <b>Chinese:</b> 好斗的, 暴躁的, 紧张的/焦虑的, 激烈的, 有暴力倾向的, 不理智的

Table 2: Korean and Chinese translation of the five mood clusters

In addition to the questions asking about music mood, the participants were asked to provide basic demographic information such as gender and age. We also asked the following two questions to determine their overall familiarity with the songs: 1) if the participants had previously heard the song, and 2) if they can name the artist and the song title. The two questions together can help gauge listeners' level of familiarity with the songs in a more objective manner than listeners' self-perception.

### 3.3 Data Analysis Methods

There were two main methods used for data analysis. First, Chi-square ( $\chi^2$ ) independence tests (Sokal & Michener, 1958) were applied to examine whether the distribution of mood judgments across mood clusters was independent from variables such as cultural background, gender, and age. The results can tell us whether each of these variables is related to mood judgments on a broad level. Chi-square distances were also calculated to further quantify the difference between two histograms (distribution of judgments across the clusters). Second, mood judgments of listeners from the same or different cultural backgrounds on the same clips were paired up to calculate the agreement ratio. This captures each listener's judgments at the finest granularity and reveals the levels of agreement reached by groups of listeners in a variety of conditions.

## 4 Results and Discussion

### 4.1 Characteristics of Participants

A total of 31 complete responses were collected from each of the three cultural groups (93 participants in total). Table 3 shows the demographics of the participants.

Cultural Background	Age			Gender	
	Min	Max	Avg.	Male	Female
American (AM)	22	55	31.8	6	25
Korean (KR)	23	39	30.7	14	17
Chinese (CN)	19	46	26.2	10	21

Table 3: Demographics of survey participants

Approximately half of the Korean and Chinese participants (13 each) responded that they stayed in the U.S. for less than a year on the date of the survey. Nine Korean and 13 Chinese participants specified 1-5

years, and nine Korean and five Chinese participants specified 6-15 years. The listeners' familiarity with the songs were categorized as follows: high familiarity means that the listener answered yes to both familiarity questions (i.e., whether they have heard the song before and whether they can identify the artist name and the song title), medium familiarity indicates yes to the first question, and low familiarity indicates no to both questions. Table 4 shows the distribution of familiarity levels across the three cultural groups.

	Low	Medium	High	No Answer	Total
AM	617	120	192	1	930
KR	832	63	33	2	930
CN	836	75	14	5	930

Table 4: Distribution of familiarity levels across three cultural groups

Overall, American listeners were much more familiar with the songs in our dataset. Korean and Chinese listeners indicated that they are not familiar with the song (i.e., no to both questions) for 89.5% and 89.9% of all responses which is significantly higher than 66.3% of American listeners. This is not surprising as the 30 clips were from the Western culture. Moreover, it is noteworthy that the Korean listeners indicated high familiarity twice as often as Chinese listeners did. This is in accordance with the fact that Korean people have been listening to Western music for a much longer period than Chinese.

## 4.2 Mood Judgments

### 4.2.1 Cultural Backgrounds vs. Mood Judgments

Figure 1 shows the distributions of mood judgments from the three groups of listeners across the five mood clusters. The proportions of songs selected for Cluster\_2 (*rollicking, cheerful, fun, sweet, amiable/good natured*) for American and Korean listeners were approximately the same and much higher than Chinese. The proportions for Cluster\_3 (*literate, poignant, wistful, bittersweet, autumnal, brooding*) for Chinese and Korean listeners were similar, and higher than Americans. Koreans were much less likely to select Cluster\_5 (*aggressive, fiery, tense/anxious, intense, volatile, visceral*) than the other two groups, and were in between of the other two groups in terms of labeling Cluster\_1 (*passionate, rousing, confident, boisterous, rowdy*) and "Other." Koreans were also more likely to select Cluster 4 (*humorous, silly, campy, quirky, whimsical, witty, wry*) than the other two groups.

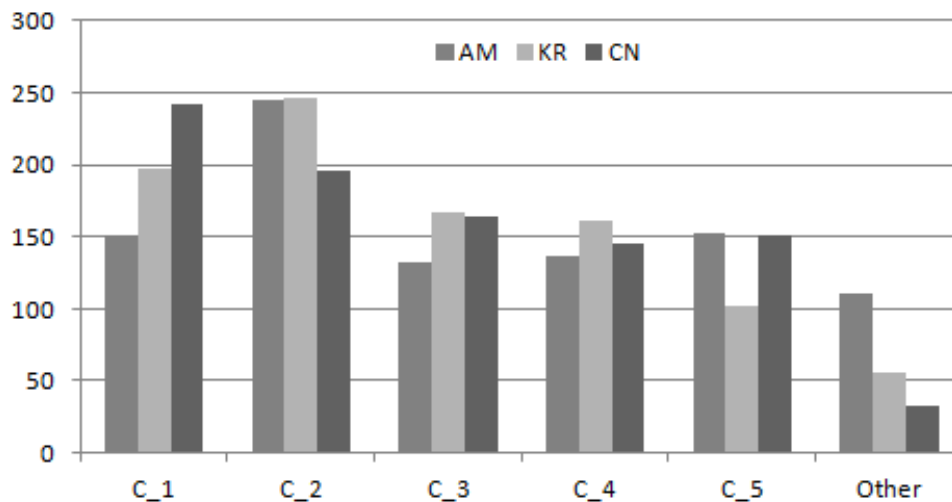


Figure 1: Mood judgments across mood clusters

The fact that American listeners tend to select “Other” more often than Korean or Chinese listeners may be due to the fact that they are more familiar with the Western songs and the cultural context in which the songs are produced; thus, their perception of the mood of a particular song maybe influenced by more than the intrinsic properties of the song. It may also reflect the differing cultural characteristics of the listeners, i.e., collectivist Eastern listeners and individualistic Western listeners as discussed in previous cross-cultural music mood research (e.g., Boer & Fischer, 2010; Hu & Lee, 2012). American listeners disagreed with the presented mood clusters approximately twice as often as Korean and three times as often as Chinese listeners.

A Chi-square test on the distributions shows that cultural background and mood judgment are not independent ( $\chi^2 = 96.72$ ,  $df = 5$ ,  $p < 0.0001$ ). In other words, there is a relationship between cultural background and mood judgments. Pair-wised tests (on each pair of two cultural backgrounds) gave the same results. To quantify the differences across cultures, Chi-square distances between each pair of mood judgment distribution (histograms) were calculated and compared. The distance between American and Chinese listeners’ judgments was the highest ( $d = 0.51$ ), followed by American and Korean listeners ( $d = 0.02$ ), and Chinese and Korean ( $d = 0.01$ ).

Overall, Chinese and Korean listeners perceived music mood more similarly, compared to American listeners. In fact, between Korean and Chinese listeners, mood judgments were significantly different for only two out of 30 songs ( $\chi^2 = 15.21 \sim 24.50$ ,  $df = 5$ ,  $p < 0.01$ ) -- both were Dance music without lyrics. American and Korean listeners significantly disagreed on nine out of 30 songs ( $\chi^2 = 9.68 \sim 26.66$ ,  $df = 4 \sim 5$ ,  $p < 0.05$ ) while American and Chinese listeners disagreed on 14 of the 30 songs ( $\chi^2 = 7.90 \sim 23.83$ ,  $df = 3 \sim 5$ ,  $p < 0.05$ ). The song with the highest disagreement between American and Korean listeners was *Got to get you into my life* by The Beatles, consistent with the result between American and Chinese listeners as reported in Hu and Lee (2012). Figure 2 shows the mood judgments on this song across the three cultural groups. Most of the American listeners selected Cluster\_2 (*rollicking, cheerful, fun, sweet, amiable/good natured*) or Cluster\_1 (*passionate, rousing, confident, boisterous, rowdy*) for this song whereas almost one-third of Korean and Chinese listeners selected Cluster\_4 (*humorous, silly, campy, quirky, whimsical, witty, wry*).

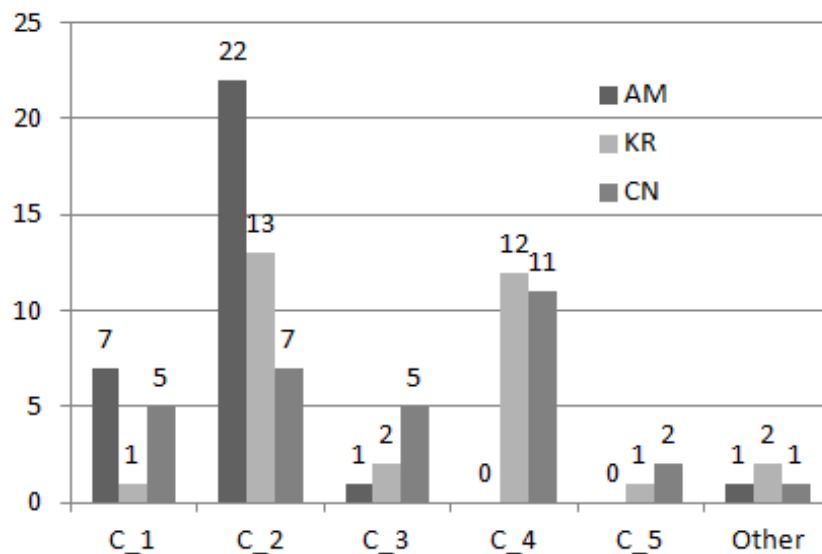


Figure 2: The mood judgments on *Got to get you into my life* by The Beatles across the three cultural groups

Based on the answers given with regard to the listeners' familiarity with the song, 29 out of 31 American listeners indicated that they had listened to this song and 19 were familiar enough with this song that they could identify the artist name and the song title. However, only four out of 31 Korean listeners had heard the song before and two could identify it, and none of the Chinese listeners had even listened to this song prior to this survey. This difference in the level of familiarity as well as the fact that Americans are much more likely to have background information on the Beatles seem to have affected how listeners determined the mood of this song.

#### 4.2.2 Agreement between Pairs of Mood Judgments

Each music clip in this survey had 31 answers from each cultural group. These answers were paired up, either within the same cultural group or across different cultural groups. Agreement ratios can be calculated by dividing the number of pairs with agreed judgments by the number of total pairs (Table 5). Korean listeners showed the lowest intra-culture agreement ratios of the three groups, indicating more diversified opinions were given by Koreans. In general, cross-cultural agreement levels were lower than intra-cultural one, except for pairs between Korean and Chinese listeners.

	American	Korean	Chinese
American	0.35	0.30	0.30
Korean		0.32	0.32
Chinese			0.35

Table 5: Agreement ratios among intra- and cross-cultural pairs of judgments

#### 4.2.3 Mood Clusters vs. Cultural Groups

Table 6 shows the number of agreement pairs across the mood clusters. It shows that different cultural groups tended to agree more on different mood clusters: Americans agreed more on Cluster\_2 (*rollicking, cheerful, fun, sweet, amiable/good natured*) and Cluster\_5 (*aggressive, fiery, tense/anxious, intense, volatile, visceral*), Koreans on Cluster\_2, and Chinese on Cluster\_1 (*passionate, rousing, confident, boisterous, rowdy*). Based on the intra-cultural agreement counts, Chi-square distances were calculated to compare each of the cultural group pairs. The results again shows the biggest difference between American and Chinese listeners ( $d = 0.5$ ), followed by American and Korean ( $d = 0.05$ ), and Korean and Chinese ( $d = 0.03$ ).

	C_1	C_2	C_3	C_4	C_5	Other	Total
AM	706	1477	778	587	1094	270	4912 (35%)
KR	1101	1332	1011	566	461	57	4528 (32%)
CN	1355	995	1203	443	894	11	4901 (35%)

Table 6: Number of agreed pairs across mood clusters among listeners within each cultural group

### 4.3 Music Characteristics

#### 4.3.1 Instrumental vs. Vocal

Half of the music clips in our dataset were instrumental, and the other half were vocal. Comparing the agreement ratios on these two types of songs can help reveal the extent to which lyrics can affect the chance

of agreement within and across cultures. Table 7 shows the comparison of intra- and cross-cultural agreement ratios for instrumental and vocal music. The lyrics of the vocal songs were in English. The fact that American listeners reached a much higher agreement ratio on vocal songs indicates that lyrics might have affected how listeners made the mood judgments. This is consistent with the findings in previous research by Lee et al. (2012) that lyrics do affect how users determine the music mood.

For the other two groups whose native language is not English, the agreement ratios on vocal music were approximately the same. As the participants were recruited over university mailing lists, we assume most of the Korean and Chinese listeners had similar levels of English fluency. However, future research is necessary to determine whether and to what extent listeners actually interpret the meaning of lyrics while listening to songs with lyrics in their second languages. In any case, vocal music seems to have helped Korean listeners reach a higher level of agreement compared to instrumental music, but not for Chinese listeners. Cross-cultural agreement ratios on instrumental music were lower than intra-cultural agreement ratios, but cross-cultural agreement ratios on vocal music were the same as intra-cultural agreement, except for the pair of AM-KR. This suggests that Korean and Chinese listeners may appreciate the mood of Western *vocal* music in similar ways, but further research is warranted to confirm this hypothesis.

	Instrumental	Vocal	All
American	0.28	0.41	0.35
Korean	0.31	0.34	0.32
Chinese	0.36	0.35	0.35
Across AM-KR	0.26	0.34	0.30
Across AM-CN	0.25	0.35	0.30
Across KR-CN	0.31	0.34	0.32

Table 7: Intra- and cross-cultural agreement ratios on instrumental and vocal music

#### 4.3.2 Genres

Analyzing the music mood judgments of the three cultural groups with respect to music genre showed that culture groups did not have a significant relationship with mood judgments. However, looking at pairs of cultural groups reveals different patterns. When considering American and Chinese participants, mood judgments and cultural groups were not independent (i.e., related) across music clips in all genres ( $\chi^2 = 21.91$ -46.68,  $df=5$ ,  $p<0.001$ ) (Hu & Lee, 2012). For American and Korean participants, mood judgments and cultural groups were independent on Pop music ( $\chi^2 = 5.22$ ,  $df = 5$ ,  $p = 0.39$ ) and Other music ( $\chi^2 = 10.09$ ,  $df = 5$ ,  $p = 0.07$ ). The Other genre included one song in each of the following genres: Folk, Metal, Reggae, Country, Oldies and Rap. This indicates that Korean and American listeners perceived the mood of Pop and Other music similarly. Even stronger similarity existed between Korean and Chinese listeners: their mood judgments and cultural background were independent on Easy-listening ( $\chi^2 = 8.85$ ,  $df = 5$ ,  $p = 0.10$ ), Rock ( $\chi^2 = 10.93$ ,  $df = 5$ ,  $p = 0.05$  and Other ( $\chi^2 = 4.73$ ,  $df = 5$ ,  $p = 0.45$ ) songs. In other words, they perceive the mood of these three genres in similar ways. Among all the genres, only Dance music always received significantly different mood judgments cross-culturally, no matter which two cultural groups were paired up.

Table 8 shows the agreement ratios within and across cultures in each genre. The numbers in bold were the most agreed upon genres. Within cultures, American listeners agreed most on Pop songs, Koreans on Easy-listening, and Chinese on Other genres. Cross-cultural comparison shows that American listeners agreed with Korean and Chinese listeners more on Pop music, while Korean and Chinese listeners agreed more on Easy-listening. The agreement on Dance music was the lowest overall. Across all genres, cross-cultural agreements were generally lower than intra-cultural agreements.



	Dance	Easy-listening	Pop	Rock	Other	All
AM	0.30	0.29	<b>0.46</b>	0.35	0.31	0.35
KR	0.25	<b>0.38</b>	0.35	0.31	0.32	0.32
CN	0.29	0.38	0.32	0.35	<b>0.41</b>	0.35
AM-KR	0.24	0.28	<b>0.36</b>	0.29	0.30	0.30
AM-CN	0.22	0.28	<b>0.33</b>	0.31	0.30	0.30
KR-CN	0.21	<b>0.38</b>	0.34	0.35	0.35	0.32

Table 8: Agreement ratios within and across cultures in each genre

## 4.4 Listener Characteristics

### 4.4.1 Gender vs. Mood Judgments

We tested the independence between the listeners' gender and distributions of mood judgments across five mood clusters. The results indicate that gender and mood judgments are not independent when aggregating judgments from all cultural backgrounds ( $\chi^2 = 36.87$ ,  $df = 5$ ,  $p < 0.0001$ ). They are also not independent when considering American listeners ( $\chi^2 = 14.99$ ,  $df = 5$ ,  $p = 0.01$ ), and Korean listeners ( $\chi^2 = 14$ ,  $df = 5$ ,  $p < 0.0001$ ), respectively. This means that male and female listeners in these two cultural groups did judge the music mood differently. However, that was not the case for Chinese listeners: male and female listeners in fact showed similar patterns in mood judgments ( $\chi^2 = 4.33$ ,  $df = 5$ ,  $p = 0.50$ ).

### 4.4.2 Age vs. Mood Judgments

The age ranges of listeners in the three cultural groups differed but all groups included listeners from 22 to 39 years old. We split this common age range into two groups, age 22 to 29, and age 30 to 39, in order to test whether mood judgments depend on age groups in the three cultures. For American listeners, the two age groups showed similar mood judgments ( $\chi^2 = 9.20$ ,  $df = 5$ ,  $p = 0.10$ ) while for Koreans and Chinese, the two age groups judged music mood differently (Koreans:  $\chi^2 = 46.89$ ,  $df = 5$ ,  $p < 0.0001$ ; Chinese:  $\chi^2 = 19.39$ ,  $df = 5$ ,  $p = 0.002$ ). This could partly be attributed to the fast changing social environments in Korea and China in the past decades. In both countries, people with age difference of about 10 years would have grown up in very different living conditions and social dynamics which certainly would have played a role in shaping a person's musical tastes. This is not the case for people who grew up in a relatively stable environment in developed countries such as the United States.

## 5 Conclusion and Future Work

Our findings reveal that the perceptions of the three cultural groups do differ. Based on the cultural groups they belong to, listeners behaved differently in their selection of mood clusters as well as agreement ratio in each of the mood clusters. However, in many aspects, Korean listeners are situated between listeners from American and Chinese cultures, evidenced by the Chi-square distances calculated on overall mood judgment distribution and pair-wise agreement on mood clusters, agreement ratio comparison between vocal and instrumental music as well as agreement ratio comparison across different music genres. Notably, Korean listeners showed the lowest intra-culture agreement ratio of the three groups.

Despite the similar level of English abilities of Korean and Chinese listeners as well as the level of familiarity to the songs in the dataset, they did perceive music moods differently in many cases. South Korea had stronger influences from Western culture over the past few decades than China. These different levels of exposure to Western music in the two countries may be one of the reasons why Korean and Chinese listeners perceived music mood differently, and why Koreans seemed to behave more similarly to Americans in much of the comparison. In addition, user characteristics affected how listeners in certain cultural groups

determine music mood: for example, gender affected American and Korean groups, and age affected Korean and Chinese groups.

The three-way comparison among American, Korean, and Chinese music listeners revealed that there was in fact some difference between Korean and Chinese listeners although they were both representing non-Western user groups. This implies that future cross-cultural music mood studies involving real users should explore similarities and differences among multiple cultural groups, not just between Western and non-Western user groups, in order to achieve a deeper understanding on cross-cultural music perception.

In our future work, we plan to conduct user interviews in order to more thoroughly investigate how listeners from different cultures perceive music mood and triangulate the results obtained from this study. In addition to the survey data we obtained from this study, it will be insightful to have people describe the reasons for selecting particular music mood in their own words. We also learned that it is difficult to find Korean and Chinese mood terms that directly map to certain English mood terms. A number of mood terms can be translated to different Korean and Chinese terms (e.g., *rollicking*, *volatile*). For some English mood terms, it can be challenging to find a single term that still preserves the original meaning (e.g., *campy*) when they are translated to Korean or Chinese. Therefore in our future work, we plan to test different translations of the mood terms to check the consistency of the results presented in this study.

## 6 References

- Balkwill, L., & Thompson, W. F. (1999). A Cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Perception*, 17(1), 43-64.
- Bischoff, K., Claudiu, S. F., Paiu, R., Nejd, W., Laurier, C., & Sordo, M. (2009). Music mood and theme classification – A hybrid approach. *Proceedings of the International Society for Music Information Retrieval Conference*, 657-662.
- Boer, D., & Fischer, R. (2010). Towards a holistic model of functions of music listening across cultures: A culturally decentred qualitative approach. *Psychology of Music*, 40(2), 179-200.
- Dang, T.-T., & Shirai, K. (2009). Machine learning approaches for mood classification of songs toward music search engine. *Proceedings of the International Conference on Knowledge and Systems Engineering*, 144-149.
- Eerola, T., Himberg, T., Toiviainen, P., & Louhivuori, J. (2006). Perceived complexity of Western and African folk melodies by Western and African listeners. *Psychology of Music*, 34(3), 337-371.
- Ellis, D., Berenzweig, A., & Whitman, B. (2003). The USPOP 2002 Pop Music Data Set. Retrieved from <http://labrosa.ee.columbia.edu/projects/musicsim/usp2002.html>
- Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A. D., & Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, 19, 573-576.
- Gregory, A. H., & Varney, N. (1996). Cross-cultural comparisons in the affective response to music. *Psychology of Music*, 24, 47-52.
- Hu, X., Downie, J. S., Laurier, C., Bay, M., & Ehmann, A. F. (2008). The 2007 MIREX Audio Mood Classification task: Lessons learned. *Proceedings of the International Society for Music Information Retrieval Conference*, 462-467.
- Hu, X., & Lee, J. H. (2012). A cross-cultural study of music mood perception between American and Chinese listeners. *Proceedings of the International Society for Music Information Retrieval Conference*, 535-540.
- Laurier, C., Sordo, M., Serrà, J., & Herrera, P. (2009). Music mood representation from social tags. *Proceedings of the International Society for Music Information Retrieval Conference*, 381-386.
- Lee, J. H., Downie, J. S., & Cunningham, S. J. (2005). Challenges in cross-cultural/multilingual music information seeking. *Proceedings of the International Society for Music Information Retrieval Conference*, 1-7.

- Lee, J. H., Hill, T., & Work, L. (2012) What does music mood mean for real users? *Proceedings of the iConference*.
- Nettamo, E., Norhamo, M., & Häkkinen, J. (2006). A cross-cultural study of mobile music: Retrieval, management and consumption. *Proceedings of OzCHI 2006*, 87-94.
- Panda, R., & Paiva, R. P. (2012). Music emotion classification: Dataset acquisition and comparative analysis. *Proceedings of the 15th International Conference on Digital Audio Effects*.
- Rana, S. A., & North, A. C. (2007). The role of music in everyday life among Pakistanis. *Music Perception*, 25(1), 59-73.
- Schäfer, T., Sedlmeier, P., & Tipandjan, A. (2008). *Music preferences in different cultures: The functions of music and their contribution to the development of music preferences in India and Germany*. Poster presented at the 25th Annual Conference of the German Society for Music Psychology.
- Serra, X., Magas, M., Benetos, E., Chudy, M., Dixon, S., Flexer, A., Gómez, E., Gouyon, F., Herrera, P., Jorda, S., Paytuvi, O., Peeters, G., Schlüter, J., Vinet, H., & Widmer, G. *Roadmap for Music Information Research*, Geoffroy Peeters (ed.).
- Sokal, R. R., & Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, 38, 1409-1438.
- Wong, P. C. M., Roy, A. K., & Margulis, E. H. (2009). Bimusicalism: The implicit dual enculturation of cognitive and affective systems. *Music Perception*, 27(2), 81-88.

## 7 Table of Figures

Figure 1: Mood judgments across mood clusters.....	263
Figure 2: The mood judgments on <i>Got to get you into my life</i> by The Beatles across the three cultural groups.....	264

## 8 Table of Tables

Table 1: Five mood clusters used in MIREX.....	261
Table 2: Korean and Chinese translation of the five mood clusters.....	262
Table 3: Demographics of survey participants.....	262
Table 4: Distribution of familiarity levels across three cultural groups.....	263
Table 5: Agreement ratios among intra- and cross-cultural pairs of judgments.....	265
Table 6: Number of agreed pairs across mood clusters among listeners within each cultural group.....	265
Table 7: Intra- and cross-cultural agreement ratios on instrumental and vocal music.....	266
Table 8: Agreement ratios within and across cultures in each genre.....	267